

【사건번호 2020-029】 국민건강보험공단 알츠하이머 치료제 등 처방환자 사망통계 데이터 사건

1. 개요

- 피신청인: 국민건강보험공단
- 대상 공공데이터: 알츠하이머 치료제 등 처방환자 사망통계
- 신청목적: 학술연구

2. 신청취지

- 신청인은 학술연구 목적으로 '특정 주성분코드에 해당하는 의약품을 처방받은 환자 중 사망한 자의 수'를 제공신청*하였으나, 피신청인이 이 사건 데이터는 상당한 가공을 통해 생산해야 하는 데이터로서 공공데이터 제공대상이 아니며, 연구용 빅데이터 서비스인 맞춤형 DB 서비스를 신청하라는 취지로 신청반려하자 분쟁조정을 신청함

* 신청인은 건강보험심사평가원으로부터 제공받은 의약품 주성분코드(9자리)별 처방환자수 데이터를 피신청인에게 제공하며 그에 대응하는 사망자수를 제공해주기를 요청하였으나, 이후 조정과정에서 '의약품 주성분코드(4자리)에 해당하는 의약품을 처방받은 환자수와 그 중 사망한 자의 수(이하 '이 사건 데이터'라 함)'로 제공범위를 조정함

3. 사실조사

가. 데이터 보유·관리 및 제공 현황

- 피신청인은 「국민건강보험법」 제14조에 따라 가입자 및 피부양자의 자격 관리, 보험급여의 관리 업무를 하고 있으며, 이와 관련하여 개인의 의약품 처방이력, 사망일 등의 데이터를 수집하여 관리하고 있음
 - 피신청인측 담당자 진술에 따르면 의약품 처방이력(처방전교부 상세정보 테이블)과 사망일(사망정보 테이블) 데이터는 별도로 관리되고 있으며,
 - 이 사건 데이터의 제공을 위해서는 '처방전 교부 상세정보 테이블'에서 특정연도에 특정 의약품을 처방받은 이력이 있는 사람들의 개인식별 아이디를 추출한 후 '사망정보 테이블'에 존재하는 개인식별 아이디와 매칭되는 사망일자 정보를 추출하여 새로운 데이터셋을 생성하여야 특정 의약품을 처방받은 자의 수 및 그 중 사망한 자의 수를 알 수 있음

나. 공공데이터 해당 여부

- “공공데이터”는 “데이터베이스, 전자화된 파일 등 공공기관이 법령 등에서 정하는 목적을 위하여 생성 또는 취득하여 관리하고 있는 광 또는 전자적 방식으로 처리된 자료 또는 정보”(공공데이터법 제2조제2호)를 의미함
- 이 사건 데이터와 관련하여 ‘환자가 의약품을 처방받은 이력 데이터’와 ‘환자의 사망일자 데이터’는 피신청인이 「국민건강보험법」에 따른 업무 수행을 통해 취득하여 관리하고 있는 정보로서, 전자적 방식으로 처리되고 있으므로 공공데이터로 볼 수 있음
- 다만 신청인이 제공신청한 데이터(특정 주성분코드에 해당하는 의약품을 처방 받은 환자 중 사망한 자의 수)를 현재 피신청기관이 생성 또는 취득하여 관리하고 있는지 여부를 판단하기 위해서는 관련 데이터의 보유·관리 현황에 관한 피신청인의 입장을 추가 청취한 후 판단할 필요가 있음

다. 공공데이터 제공대상 해당 여부

- 공공기관은 정보공개법상 비공개대상정보 또는 제3자의 권리정보가 포함되어 있는 등의 예외사유를 제외하고는 보유·관리하고 있는 공공데이터를 국민에게 제공하여야 함(공공데이터법 제17조제1항)
- 단, 공공기관은 보유·관리하고 있는 데이터를 제공할 의무를 가질 뿐, 요청에 따라 추가적으로 공공데이터를 생성하거나 변형 또는 가공, 요약, 발췌할 의무를 부담하지 않음(공공데이터법 제26조제3항)
- 이와 관련하여, 정보공개법상 공공기관이 보유·관리하는 정보의 의미 및 새로운 가공의 의미에 대해 대법원은 “그 정보가 청구인이 구하는 대로는 되어 있지 않다고 하더라도 공개청구를 받은 공공기관이 공개청구대상 정보의 기초자료를 전자적 형태로 보유·관리하고 있고, 당해 기관에서 통상 사용되는 컴퓨터 하드웨어 및 소프트웨어와 기술적 전문지식을 사용하여 그 기초자료를 검색하여 청구인이 구하는 대로 편집할 수 있으며, 그러한 작업이 당해 기관의 컴퓨터 시스템 운용에 별다른 지장을 초래하지 아니한다면, 그 공공기관이 공개청구대상정보를 보유·관리하고 있는 것으로 볼 수 있고, 이러한 경우에 기초자료를 검색·편집하는 것은 새로운 정보의 생산 또는 가공에 해당한다고 할 수 없다(대법원 2010.2.11., 선고 2009두6001 판결)”고 판시한 바 있음

- 위원회는 국민의 데이터 이용 편의를 위하여 공공데이터의 가공 등이 용이한 경우*나, 가공에 소요되는 비용을 신청인이 부담하는 경우** 등 공공데이터의 가공 등에 공공기관의 비용 또는 노력이 과도하게 투입되지 않을 것을 전제로 제공을 적극 검토하도록 권고한 바 있으며,
 - * 행정자치부 봉사참여서비스 오픈API의 검색서비스 결과값에 주소코드가 포함되도록 수정해 줄 것을 요청한 사안(분쟁조정 2016-015사건)에서, 피신청인에게 이용자의 요청에 따라 공공데이터를 추가적으로 수정·제공할 의무가 없는 것이 원칙(공공데이터법 제26조제3항)이나, 기술적 검토 결과 신청인이 원하는 형태로의 수정이 용이한 점 및 해당 오픈API의 활용도 제고가 필요한 점을 고려하여 수정·제공토록 조정(사전조정)
 - ** 국토교통부 자동차등록정보(2015-012, 2017-028), 소방청 화재발생지 상세주소데이터(2018-012)의 경우 데이터의 가공 및 치환비용을 신청인이 부담하는 것을 조건으로 데이터를 제공토록 조정
- 데이터의 제공을 위해 별도의 가공(데이터 추출 및 오류 정제 등)과정이 필요하고 해당 가공을 위해 상당한 시간 및 노력이 소요될 것으로 보이는 경우에는 제공의무를 인정하기 어렵다고 본 사례가 있음(대검찰청 범죄분석 데이터 사건(2018-020))
- 피신청기관 제출자료 및 담당자 진술에 따르면 이 사건 데이터가 별도의 DB에 존재하며 신청인이 원하는 형태로 관리되고 있지는 않은 것으로 보이나, 이 사건 데이터의 제공을 위한 기초자료를 보유하고 있는 것은 분명하므로 피신청인이 보유하고 있는 하드웨어·소프트웨어 및 전문지식을 사용하여 신청 데이터를 추출할 수 있고, 그와 같은 작업이 피신청인의 시스템 운용 등 업무에 지장을 초래하지 않을 경우 제공을 적극적으로 검토하도록 할 필요가 있음

4. 조정내용

가. 조정결정 사항

- o 피신청인은 이 사건 데이터*를 '21.2.28.까지 신청인에게 제공한다.
 - * 데이터 제공범위에 관한 상세내용은 별지 참조
- 신청인은 이 사건 데이터를 활용한 연구결과를 공표할 때, 데이터의 출처 및 산출조건과 함께 피신청인은 연구에 관여하지 않았으며 데이터 분석 및 연구결과는 온전히 신청인의 판단에 의한 것임을 명시하여야 한다.

나. 조정결정 이유

- 「공공데이터의 제공 및 이용 활성화에 관한 법률(이하 ‘법’이라 함)」에 따르면, 공공기관은 보유·관리하고 있는 공공데이터를 국민에게 제공하여야 함(법 제17조제1항)
 - 다만, 「공공기관의 정보공개에 관한 법률」 제9조에 따른 비공개대상정보 또는 「저작권법」 및 그 밖의 다른 법령에서 보호하고 있는 제3자의 권리가 포함된 것으로 해당 법령에 따른 정당한 이용허락을 받지 않은 정보를 포함하는 경우는 제외됨(법 제17조제1항각호)

- 이 사건 관련 법령, 피신청인 제출자료 및 진술 등을 고려하면, 이 사건 데이터는 피신청인이 보유·관리하고 있는 공공데이터로서 공공데이터법 제17조제1항에 따른 제공대상 데이터에 해당한다고 판단됨
 - 피신청인은 내부 심의위원회가 신청인 연구의 공익적 목적 등을 심의한 후 데이터 제공을 승인하기 전에는 이 사건 데이터를 제공하기 어렵다고 주장하나, 이 사건 데이터가 공공데이터법상 제공대상 데이터에 해당하는 이상 피신청인은 해당 데이터를 제공할 의무가 있으므로, 내부 절차를 거치지 않았다는 점은 공공데이터 제공거부사유로 볼 수 없음
 - 다만, 피신청인이 신청인의 연구결과가 피신청인의 공식입장으로 오해되거나 그로 인한 예상치 못한 문제의 발생을 우려할 수 있으므로, 신청인이 이 사건 데이터를 활용한 결과를 공표할 때에는 데이터의 출처 및 산출 조건과 함께, 피신청인이 연구에 관여하지 않았으며 데이터의 분석 및 연구 결과는 온전히 신청인의 판단에 의한 것이었음을 명시하도록 함

5. 조정결과

- 조정성립

[별지] 공공데이터제공분쟁조정위원회 2020-029사건 데이터 제공범위

1. 이 사건 데이터는 신청인이 제출한 주성분코드(앞의 4자리)에 해당하는 의약품을 처방받은 환자 수(이하 '처방자수'라 함) 및 해당 환자 중 사망한 자(이하 '사망자수'라 함)의 수로 정의함
2. 상기 주성분코드(앞의 4자리) 중 1486, 1900, 2245, 1615, 2040, 2242, 2270, 3786, 4515, 4748의 경우 각각의 코드에 대한 데이터를 제공함
 - 2-1. 3852, 6434, 1625, 1679, 1680, 2063, 2429(총 7개)는 그룹화한 후(하나의 군으로 묶음), 해당 그룹에 대한 데이터를 제공함
3. 처방(진료) 시점은 2010.1.~2019.12.를 기준으로 하고 연도별로 처방자수를 산출함
 - 3-1. 사망시점은 데이터 산출시를 기준으로 함
4. 이 사건 데이터는 신청인이 공공데이터 제공신청 시 피신청인에게 제출한 건강보험심사평가원 데이터와 연계되지 않으며, 피신청인이 보유·관리하는 데이터를 기준으로 산출함
5. 이에 따른 제공데이터의 예시는 다음의 표와 같음

연도(처방)	주성분코드	처방자수	사망자수
2010년	1486		
	1900		
	2245		
	1615		
	2040		
	2242		
	2270		
	3786		
	4515		
	4748		
	3852		
	6434		
	1625		
	1679		
	1680		
	2063		
	2429		